

논문 2010-47CI-2-4

# RSPM : NVR 시스템 기반의 저장장치 신뢰성 향상 기법

## ( RSPM : Storage Reliability Scheme for Network Video Recorder System )

이 근 형\*, 송 재 석\*, 김 덕 환\*\*

( Geunhyung Lee, Jae-Seok Song, and Deok-Hwan Kim )

### 요 약

NVR은 시스템을 구성하는 모든 카메라와 비디오 서버가 네트워크 환경으로 연결되어 설치가 용이하며 유지 및 보수의 효율성이 뛰어나 차세대 영상 감시 시스템으로 각광받고 있다. 또한 사용자가 웹 기반으로 카메라, 비디오 서버에 접근 할 수 있어 시간적, 공간적 제약을 극복한 관리환경을 제공한다. 하지만 기존의 NVR 저장장치는 데이터가 손상될 경우, 이를 복구할 방법과 실시간으로 시스템을 유지할 수 있는 수단이 부족하다. 본 논문에서는 데이터의 손상 복구 및 실시간성 유지가 가능한 RSPM(Reliable Storage using Parity and Mirroring)기법을 제안한다. RSPM기법은 패리티 코드와 Liberation 코드 기법을 적용하여 영상 데이터의 손상을 복구할 수 있으며 동적 미러링 기법을 적용하여 시스템 데이터의 손상 복구 및 실시간성을 유지할 수 있다. 실험을 통해 Liberation 코드를 이용한 RSPM 기법이 EXT3 파일 시스템보다 영상 데이터 손상에 대한 손실율이 평균 약 11.29%의 감소됨을 알 수 있으며 패리티 코드를 적용한 RSPM 기법 보다 약 5.21% 감소됨을 알 수 있다.

### Abstract

Network Video Recorder becomes popular as a next generation surveillance system connecting all cameras and video server in network environment because it can provide ease of installation and efficient management and maintenance. But in case of data damage, the storage device in traditional NVR has no recovery scheme and it is disabled in processing real-time requests. In this paper, we propose an Reliable Storage using Parity and Mirroring scheme for improving reliability on storage device and maintaining system on realtime. RSPM uses a Liberation coding to recover damaged multimedia data and dynamic mirroring to repair corrupted system data and to maintain real-time operation. RSPM using the Liberation code is 11.29% lesser than traditional file system and 5.21% less than RSPM using parity code in terms of loss rate of damaged multimedia data.

**Keywords :** NVR, HDD, RSPM, channel code, dynamic mirroring

### I. 서 론

최근 촬영 기술의 발달과 네트워크 기술의 발전 및 비용 하락에 의하여 영상 감시 시스템은 기존에 비하여

더욱 폭넓게 사용되고 있다. 일반적으로 사용되는 보안 용도뿐만 아닌 초등학교 등하교길 감시, 공장 생산 라인 감시, 쇼핑몰 손님 현황 분석 등, 다양한 분야에서 활용되고 있으며 근래 지속적인 발전을 보이고 있다. 과거 아날로그 방식으로 영상을 촬영하여 비디오테이프에 기록하는 VCR(Video Cassette Recorder) 방식은 기록 매체를 비디오테이프에서 디지털 미디어로 변경하는 DVR(Digital Video Recorder)시스템으로 발전하였다. 또한 DVR은 네트워크 기술의 발전 및 인프라가 구축됨에 따라 그림 1과 같이 IP 카메라와 비디오서버에 원

\* 학생회원, \*\* 정회원, 인하대학교 전자공학과  
(Department of Electronic Engineering,  
Inha University)

※ 이 논문은 인하대학교 연구비 지원에 의하여 연구되었음.

접수일자: 2009년12월16일, 수정완료일: 2010년3월10일

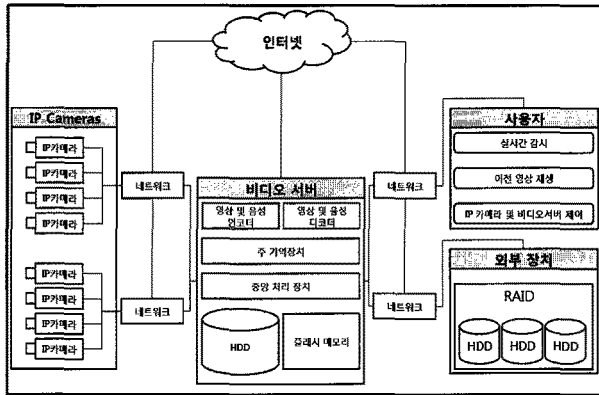


그림 1. Network Video Recorder의 개요  
Fig. 1. Overview of Network Video Recorder.

적으로 접속하여 재생 및 제어를 수행할 수 있는 NVR(Network Video Recorder) 시스템으로 거듭나게 된다.

기존의 NVR 시스템은 전체 시스템의 규모에 따라 수량의 차이는 있지만 공통적으로 하드 디스크를 저장 장치로 사용하였다. 또한 대체적으로 개인용 PC 환경에서 사용하는 것과 동일한 파일 시스템을 사용하였다. 하지만 개인용 PC 환경의 파일 입출력 요청과 NVR 시스템 환경의 파일 입출력 요청은 서로 다른 특성을 보이고 있다. 개인용 PC 환경의 파일 입출력 요청은 대부분 그 크기가 작으며 공통적인 패턴 없이 무작위 발생하는 특성을 보이고 있지만 NVR 시스템의 파일 입출력 요청은 각 요청의 크기가 크며, 일정한 패턴이 반복되는 특성을 보이고 있다.<sup>[8]</sup> 따라서 이러한 특성을 활용하여 하드 디스크의 입출력 성능을 최대한 높일 수 있는 파일 시스템이 필요하다.

NVR 시스템은 감시, 촬영한 영상 데이터를 저장장치에 저장, 관리하여 사용자가 이에 접근할 수 있도록 하며 저장된 영상 데이터를 사고나 범죄 발생 시 원인 분석의 중요한 단서로 활용된다. 따라서 높은 신뢰도가 보장되어야 하며 사용자도 구매 시 신뢰도를 가장 먼저 고려한다. 하지만 NVR 시스템은 그 역할상 지속적으로 저장장치에 대용량의 쓰기 연산을 수행하게 되어 저장장치의 노화가 우려된다. 또한 이에 따라 불량 섹터가 많아지며 데이터의 손상 가능성이 높아지고 불안정한 네트워크 상황, 사용자의 잘못된 조작, 외부 충격 등에 의해서 촬영된 영상 데이터 및 시스템 데이터가 손실될 수 있다. 하지만 기존 NVR에 사용되는 파일 시스템은 이러한 손실에 충분히 대비되어 있지 않다.

본 논문은 NVR 시스템에서 충분한 저장장치의 신뢰

성을 확보하기 위하여 RSPM(Reliable Storage using Parity and Mirroring)기법을 제안한다. 제안하는 기법은 XOR 연산을 사용해 손상 데이터 복구용 부호어를 생성하는 Liberation 코드를 적용하여 손상된 영상 데이터를 복구한다. 또한 시스템 데이터 손상 시, 이를 즉시 대체할 수 있도록 사본을 보관하는 미러링 기법을 이용하여 시스템 데이터의 손상 복구와 NVR 시스템의 실시간성을 유지한다. 그리고 영상 데이터의 연속적 쓰기 기법을 이용하여 하드 디스크의 읽기/쓰기 성능을 극대화한다.

본 논문의 순서는 먼저 II장에서 관련 연구에 대해 소개하고 III장에서 RSPM 기법의 설계 및 구조에 대해 설명한다. IV장에서는 RSPM 기법에 적용된 채널 코딩 기법 및 미러링 기법을 설명하였으며, V장에서 본 기법에 대한 평가를 서술하였다. 마지막으로 VI장에서 결론을 맺는다.

## II. 관련 연구

### 1. 하드 디스크의 오류 원인

저장장치는 다양한 이유로 오류를 일으킨다. 오류의 종류는 크게 데이터의 분실과 손실로 분류할 수 있다.<sup>[4-5]</sup>

데이터의 분실은 Servo영역 손실, 헤드의 트랙 이탈, SMART 기한 초과, 컨트롤러 불량 등의 원인에 의해 일어나며 주로 하드웨어 또는 펌웨어 영역에서 발생한다. 데이터의 분실 오류는 오류가 일어나는 영역이 크고 명확하여 오류의 탐지가 용이하다. 따라서 많은 연구가 이루어졌으며 충분한 대책이 이루어졌다.

데이터의 손실은 장치노화, 비트 전송 오류, 과열 또는 부식에 의한 데이터 손실, 플래터의 흠집에 의한 손실 등에 의하여 발생되며 오류가 국지적으로 일어난 섹터에 실제 데이터의 입출력이 시도되지 않으면 탐지되지 않는 잠재적인 오류이다. 대표적인 데이터 손실 오류는 하드 디스크의 불량 섹터 오류로 하드 디스크 특정 섹터의 데이터를 완전히 소실할 수 있다.

### 2. 복수의 저장장치를 이용한 신뢰성 향상

데이터의 분실 오류를 해결하기 위해 일반적으로 복수의 하드 디스크를 활용하는 RAID(Redundant Arrays of Inexpensive Disks)기법을 적용한다<sup>[3]</sup>. RAID 기법은 복수의 하드 디스크에 분산하여 데이터를 저장하여 입

출력 성능을 향상시키는 스트라이핑 기법, 복사본을 만들어 두는 미러링 기법, 비트 단위의 오류를 검출, 복구 하는데 적합한 ECC(Error Correction Code), 블록 단위의 오류를 검출, 복구하기 위한 패리티 기법 등을 이용하여 성능과 신뢰성을 향상하였다.<sup>[11]</sup> 하지만 규모가 작은 시스템을 구성할 시, 복수의 저장장치를 사용하기 어려운 경우가 있다. 또한 RAID 기법은 일반적으로 디바이스 드라이버 또는 펌웨어 영역에서 구현되기 때문에 데이터의 특성에 따라 세분화 하여 신뢰성 기법을 적용할 수 없다.

### 3. EXT3

많은 NVR 시스템이 임베디드 리눅스 환경에서 작동하고 있다. 따라서 리눅스 환경의 범용 파일 시스템이라 할 수 있는 EXT3(Extended File System 3)가 많이 사용되고 있다. EXT3는 EXT2가 발전된 형태로 그림 2와 같은 저널링 기능이 추가되어 신뢰성을 향상시켰다.

저널링은 파일 시스템의 메타데이터(U), inode(Ni)와 데이터 블록(Nd)의 쓰기 도중 오류가 발생되어 파일 시스템 전체에 악영향을 미치는 것에 대비한 것으로, 데이터가 실제로 쓰기 전에 먼저 저널에 쓰게될 데이터를 복사 해두고 실제 쓰기를 수행하는 것을 말한다. 만약 저널에 쓰기 도중 문제가 발생할 경우 쓰기 요청 자체가 삭제되어 파일 시스템의 무결성을 보존할 것이다. 또한 디스크에 직접 쓰기 도중 문제가 발생할 경우 문제 복구 후 다시 저널에서 디스크로 쓰기 시도를 할 것이다.

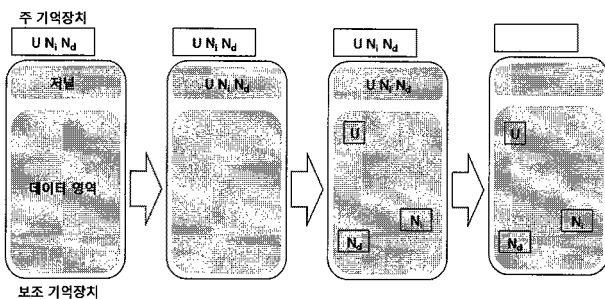


그림 2. EXT3 저널링 기법  
Fig. 2. EXT3 Journaling.

## III. RSPM 기법

### 1. RSPM 파일시스템 구조

NVR 시스템의 비디오 서버는 기본적으로 하나의 하드 디스크를 가진다. 하드 디스크 내의 불량 섹터가 발생하여 데이터의 일부가 손상되는 경우가 종종 발생한다. 이에 대한 대응책이 기존의 NVR 시스템에는 마련되어 있지 않다. 이를 극복하기 위하여 하나의 플래시 메모리를 추가 장착하는 RSPM 기법을 제안한다. 추가된 플래시 메모리에 시스템 데이터의 사본, 채널 코딩 기법의 부호어 일부를 저장한다. 이로써 하드 디스크에 저장된 데이터들의 손상을 복구할 수 있다,

그림 3은 RSPM 기법이 적용된 파일 입출력 계층의 개요를 나타낸다. 응용프로그램 계층은 사용자에게 의해 작성된 응용 프로그램을 나타내며 시스템 콜을 통해 가상 파일 시스템(Virtual File System, VFS)에 접근한다. 가상 파일 시스템은 RSPM 파일 시스템과 상호작용하여 RSPM 블록 디바이스 드라이버에 적합한 입출력 요청을 생성한다. RSPM 파일 시스템은 EXT3 파일 시스템을 기반으로 입출력 요청에 패리티 코드나 Liberation 코드와 같은 채널 코딩<sup>[2]</sup>과 동적 미러링 기법을 적용한다. RSPM 블록 디바이스 드라이버는 하나의 하드 디스크와 플래시 메모리를 하나의 단위 디바이스로 간주한 RSPM 블록 디바이스를 제어한다.

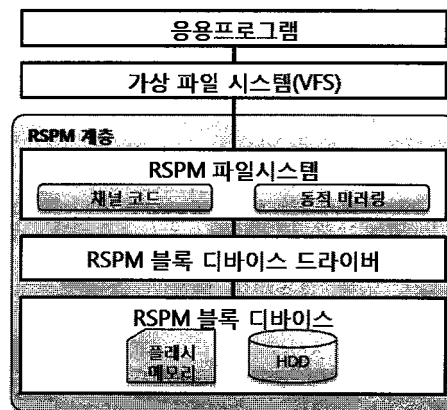


그림 3. RSPM 파일시스템 개요  
Fig. 3. The overview of RSPM file system.

### 2. RSPM 블록 디바이스의 신뢰성 측정

저장장치의 신뢰성 단위는 MTTF(Mean Time To Failure)를 사용한다. 이는 고장과 다음 고장 사이의 평균 시간을 말하며 제품의 데이터 시트에 일반적으로 포함되어 있다. MTDDL(Mean Time To Data Loss)는 복수개의 저장장치를 이용하여 신뢰성을 확보하는 기법에서 등장한 단위로, 실제 데이터의 손실과 다음 데이

터의 손실 사이의 평균 시간을 말한다.

모든 하드 디스크와 플래시 메모리는 하드웨어 스펙에 MTTF를 포함하고 있으며 RSPM 기법은 플래시 메모리와 하드 디스크를 동시에 사용하므로 MTDDL을 계산할 수 있다<sup>[6]</sup>.

$$MTDDL_{RSPM} = \frac{MTTF_{HDD} * MTTF_{FLASH}}{MTTF_{HDD} + MTTF_{FLASH}} + \frac{1}{P_{rRSPM}} \quad (1)$$

$P_{rRSPM}$ 은 데이터 복구 중 데이터의 손실이 일어나는 경우로 두 가지 경우에 발생할 수 있다. 플래시 메모리가 오류를 일으킨 후 하드디스크가 오류를 일으킨 경우와 하드디스크가 오류를 일으킨 후 플래시 메모리에 오류가 발생한 경우이다. 따라서 데이터 손실 가능성은 다음과 같다.

$$P_{rRSPM} = P_{rFLASH} * P_h + (1 - P_{rFLASH}) * P_f \quad (2)$$

$P_{rFLASH}$ 는 두 개의 장치 중 플래시 메모리가 먼저 오류를 일으킬 가능성이다.  $P_h$ 은 데이터 복구 중 하드 디스크가 오류를 일으킬 가능성이며  $P_f$ 는 플래시 메모리가 오류를 일으킬 가능성이다. 플래시 메모리의 오류를 복구하는데 걸리는 시간을  $t_{rFLASH}$ 라고 정의하면 지수 분포에 따른 오류율에 의해  $P_h = 1 - e^{-t_{rFLASH}/MTTF_{HDD}}$ 로 가정한다.  $t_{rFLASH} \ll MTTF_{HDD}$ 이므로 데이터 복구 중 하드 디스크가 오류를 일으킬 가능성  $P_h \approx \frac{t_{rFLASH}}{MTTF_{HDD}}$

이다. 이와 비슷하게  $P_f \approx \frac{t_{rHDD}}{MTTF_{FLASH}}$ 이다.

이를 수식 (2)에 적용하면 다음과 같다.

$$\begin{aligned} Pr_{RSPM} &= \frac{MTTF_{HDD}}{MTTF_{HDD} + MTTF_{FLASH}} \frac{t_{rFLASH}}{MTTF_{HDD}} \\ &+ \frac{MTTF_{FLASH}}{MTTF_{HDD} + MTTF_{FLASH}} \frac{t_{rHDD}}{MTTF_{FLASH}} \quad (3) \\ &= \frac{t_{rFLASH} + t_{rHDD}}{MTTF_{HDD} + MTTF_{FLASH}} \end{aligned}$$

이를 수식 (1)에 적용하면 다음과 같은 결론이 나온다.

$$\therefore MTDDL_{RSPM} = \frac{MTTF_{HDD} * MTTF_{FLASH}}{t_{rHDD} + t_{rFLASH}} \quad (4)$$

표 1은 수식 (4)를 적용하여 기존의 단일 디스크와 RSPM 기법을 적용하였을 경우의 MTDDL을 계산한 결과이다.

표 1. 단일 하드 디스크와 RSPM 적용에 따른 MTDDL 비교 ( $MTTF_{HDD} = 300000$  시간,  $MTTF_{FLASH} = 50000$  시간)

Table 1. Comparison of MTDDL ( $MTTF_{HDD} = 300000$  hours,  $MTTF_{FLASH} = 50000$  hours).

	MTDDL(년)
단일 디스크	5.7
RSPM	5265097

### 3. 데이터 타입에 따라 신뢰성 기법 차별 적용

일반적인 PC 환경 저장장치의 파일들은 제각각 그 역할과 특성이 다르며, 이를 명확하게 구분 지을 수 없어 각 파일의 특성에 적합한 처리를 할 수 없다. 하지만 NVR 시스템은 영상 및 음성을 나타내는 미디어 데이터와 시스템을 유지하는데 필요한 시스템 데이터, 크게 두 가지로 분류할 수 있다.

표 2는 각 데이터 타입에 따른 신뢰성 기법을 보여준다. 미디어 타입은 카메라에서 촬영된 영상 또는 사진, 음성을 뜻하며 NVR 시스템의 핵심적인 데이터이므로 높은 신뢰성을 확보할 필요가 있다. 또한 미디어 타입은 크기가 크며 프레임 복원(특정 프레임이 손상 시, 인접한 프레임과 비교하여 손상된 프레임을 복원)과 같은 2차적인 복원 수단이 존재하기 때문에 높은 복원을 보다 공간 효율 및 과부하 방지를 위해 작은 크기의 부호어를 생성하는 패리티 코드나 Liberation 코드와 같은 채널 코딩 기법을 적용한다. 채널 코딩기법으로 인코딩된 부호어는 플래시 메모리에 저장 되어 하드디스크의 데이터 블록들이 손실 되었을 경우 이를 복구하기 위해 사용된다.

시스템 타입은 NVR 시스템의 운영 체제 또는 전용 응용프로그램에서 사용하는 파일들이며 NVR 시스템의 정상 가동을 위해서 필수 불가결한 파일들이다. 따라서 손상 시 즉시 복구할 수 있도록 복사본을 플래시 메모리에 작성하는 미러링 기법을 적용한다.

그림 4는 데이터의 분류 및 시스템 타입 데이터의 쓰

표 2. 데이터 타입에 따른 적용 신뢰성 기법  
Table 2. Reliability policy for each data type.

데이터 타입	크기	실시간성 유지	채널 코드	미러링
미디어	대	-	O	-
시스템	소	O	-	O

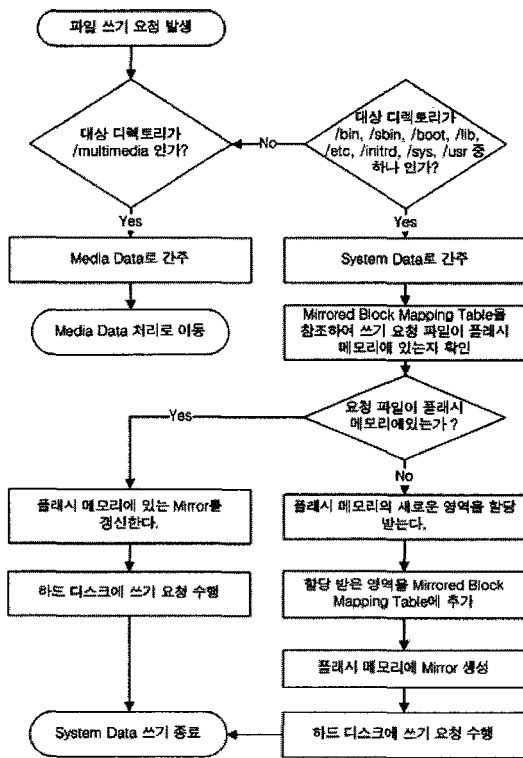


그림 4. 데이터 타입의 분류 및 시스템 타입 처리 순서도  
 Fig. 4. A Flow chart of classification of Data type and processing of write operation for System type.

기 연산에 대한 순서도이다. RSPM 파일 시스템은 파일이 저장된 디렉토리에 따라 타입을 구분하며 이는 파일 시스템 설계자에 의해 수정될 수 있다. 따라서 카메라 데이터를 저장하는 응용 프로그램은 이에 맞게 데이터를 배치할 수 있어야 한다. /bin, /sbin, /lib, /etc, /boot, /initrd, /sts, /usr 디렉토리에 쓰여지는 데이터는 시스템 타입으로 분류된다. 시스템 타입의 쓰기 처리는 미러링 기법에 따라 플래시 메모리에도 동시에 업데이트 하게 된다. 미디어 타입은 /multimedia에 쓰여지는 파일이다. 미디어 타입의 쓰기 연산 처리는 다음절에서 설명한다.

4. 미디어 타입을 위한 쓰기 기법

NVR 시스템은 하드 디스크를 저장장치로 활용한다. 하드 디스크는 회전하는 플래터에 암을 이용하여 데이터를 읽으며, 이러한 기계적인 특성에 의해 불연속적인 읽기/쓰기의 성능이 좋지 않으며 상대적으로 연속적인 읽기/쓰기에 좋은 성능을 보인다. 따라서 비디오 서버는 이러한 하드 디스크의 특성을 고려하여 연속적인 읽기/쓰기를 수행할 수 있도록 유도하여야 한다.

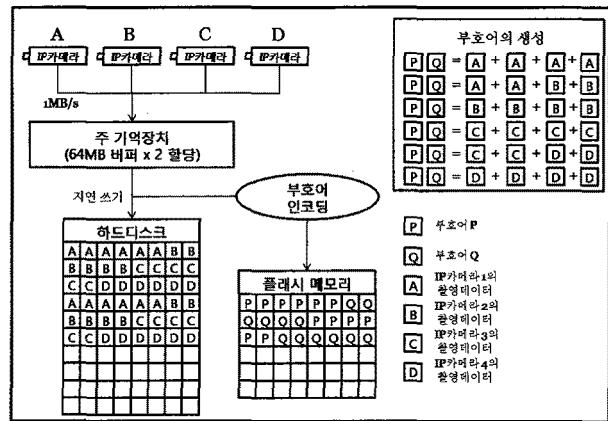


그림 5. 자연쓰기 기법의 예제  
 Fig. 5. An example of Write-back scheme.

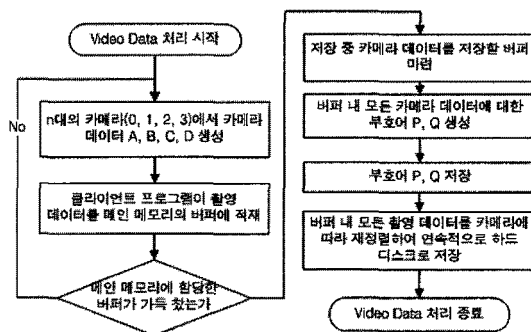


그림 6. 자연쓰기 기법의 순서도  
 Fig. 6. A Flow-chart of Write-back.

제안하는 기법은 그림 5와 같이 각 채널 당 미디어 타입 데이터가 연속적으로 쓰이게 하며, 부호어를 효율적으로 인코딩하기 위해 자연쓰기 기법을 적용한다. 자연쓰기는 메인 메모리를 버퍼로 활용하여 미디어 타입 데이터를 일정량 모으고 이를 방출할 시 일정 단위로 묶어 부호어를 생성한다. 버퍼에 모인 카메라 데이터는 채널 번호를 기준으로 정렬되어 방출되며 이로써 공간 지역성을 유지할 수 있다. 자연쓰기의 부호어는 채널에 상관없이 연속적으로 쓰인 데이터를 이용하여 생성한다. 여기서 부호어 P와 Q는 플래시 메모리에 저장한다. 자연쓰기의 쓰기 연산은 그림 6과 같다. 자연쓰기는 각 채널에서 카메라 데이터 A, B, C, D 를 모아 버퍼가 가득 찰 때 까지 적재한다. 버퍼는 이중으로 구성되며 버퍼가 방출 될 시에도 다른 버퍼에 적재가 가능하도록 구성한다. 버퍼가 가득 차면 부호어 P와 Q를 생성하여 플래시 메모리에 저장한다. 부호어 P, Q의 쓰기 연산이 끝난 후에 실제 미디어 데이터의 저장이 이루어진다.

### IV. 신뢰성 향상 기법

#### 1. Liberation 코드

Liberation 코드<sup>[1]</sup>는 RAID-6 기반 채널 코딩 기법 중 하나로 스트라이핑 된 데이터의 XOR 연산으로 부호어를 만들어 데이터 손상 시 복구하는 기법이다. Liberation 코드는 현재 제안되어진 채널 코딩 기법 중 인코딩과 디코딩 연산에서 수행하는 XOR 연산의 수가 적어 높은 성능을 보이고 있다. 따라서 RSPM 기법은 Liberation 코드를 채용한다.

#### 가. 인코딩

Liberation 코드는 그림 7과 같은 이진 분산 행렬(Binary Distribution Matrix, BDM) 기반 코드로 부호어(Codeword) P, Q를 작성한다. 각 벡터는 w 개의 비트로 이루어져 있으며, 각 매트릭스는  $w \times w$ 개의 비트로 이루어져 있다. BDM의 I 행렬은 항등원 행렬(Identity Matrix)로 XOR 곱셈 연산을 하여도 같은 결과를 내는 행렬이다.

BDM의  $X_p$ 와  $X_q$ 는 부호어를 생성하는 부호 분산 행렬(Coding Distribution Matrix, CDM)로 P와 Q를 생성하는데 사용된다. P는 각 스트라이프의 XOR 연산으로 정의되며 이는 수식 (5)와 같다.

$$P = D_0 \oplus D_1 \oplus \dots \oplus D_{k-1} \tag{5}$$

부호어 P는 플래시 메모리에 저장된다.

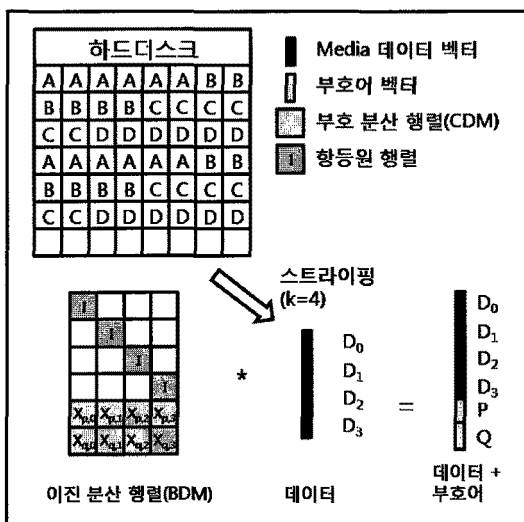


그림 7. 비트 행렬 코딩 시스템의 인코딩 연산 예제  
Fig. 7. An Example of Encoding of Bit-matrix Coding System.

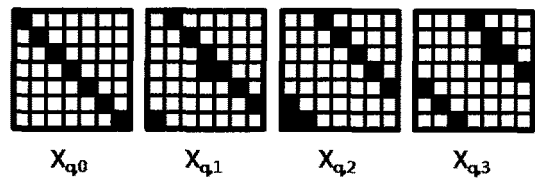


그림 8.  $X_q$ 의 예제( $w = 7, k = 4$ )  
Fig. 8. An Example of  $X_q$  ( $w = 7, k = 4$ ).

부호어 Q를 계산하기에 앞서 그림 7과 같은 Liberation 코드 고유의 CDM을 생성할 필요가 있다. CDM은 다음과 같은 법칙에 의해 생성된다.

$$\begin{aligned}
 & - X_{q,0} = I \\
 & - 0 < i < k, X_{q,i} = L_{\rightarrow i} + O_{y,y+i-1} \\
 & \quad y = \frac{i(w-1)}{2}
 \end{aligned}$$

그림 8의 첫 번째 행렬  $X_{q,0}$ 는 항등원 행렬이며,  $L_{\rightarrow i}$ 는  $i$ 가 증가함에 따라 행렬값이 오른쪽으로 1 비트씩 이동하는 것을 타나낸다. 또한 추가적으로  $O_{y,y+i-1}$  비트를 1로 설정한다.

예를 들어  $D_0 = 1001101, D_1 = 0010110, D_2 = 1011001, D_3 = 0110111$  로 각 스트라이프를 가정하면, 각 스트라이프 간 XOR 연산을 통해  $P = 0110101$ 이 된다. 또한 D와  $X_q$ 를 XOR 곱셈 연산하여  $Q = 0101101$ 이 됨을 알 수 있다.

#### 나. 디코딩

만약 데이터가 손상되었을 경우, 부호어를 이용하여 이를 복구할 수 있어야한다. 이러한 연산을 디코딩 연산이라 하며 그림 7의 연산을 역으로 수행하여 원본 데이터를 복구한다. (손상되지 않은 데이터(생존자, Survivor) + 부호어) \* BDM 중 손상된 영역에 해당하는 부분을 삭제한 행렬의 역원(Inverse Matrix) 중 손상 영역에 해당되는 행렬을 연산하여 손상된 영역을 복원한다.

#### 2. 동적 미러링

RSPM 기법의 미러링 기법은 파일 시스템의 블록 단위로 작동하며 저장장치 영역에서 일반적으로 사용되는 미러링 기법과 다르게 사본을 동적으로 할당하여 보관하고자 한다.

최근 플래시 메모리의 발전에 따라 저장 공간은 지속적으로 확장되었다. 하지만 비디오 서버를 구동시키는

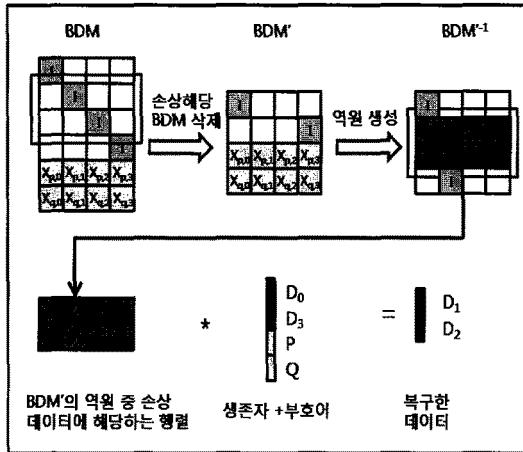


그림 9. 비트 행렬 코딩 시스템의 디코딩 연산 예제  
 Fig. 9. An example of decoding of Bit-matrix Coding System.

임베디드 기반 운영체제의 필요 저장 공간은 크지 않으며 앞으로 그 증가폭도 크지 않으리라 예상된다. 기존의 미러링 방식은 하드웨어에 의한 장치와 장치의 1:1 매핑 방식이며 이를 채용할 경우, 시스템 데이터가 포함된 저장장치와 같은 용량을 가지고 있는 저장장치를 사용하여야 한다.

하지만 사본의 저장 영역을 동적으로 할당하면 사본이 저장되는 영역 이외의 영역을 활용할 수 있으며 운영체제가 설치된 저장장치와 다른 종류의 저장장치를 활용할 수 있어 유연하다.

동적 미러링 기법을 위해 미러 블록 매핑 테이블 (Mirror Block Mapping Table, MBMT)을 제안한다. MBMT는 사본이 저장될 플래시 메모리가 마운트 될 시 생성되며, 원본 시스템 데이터 블록 갱신 시 사본에 변경 내용을 적용 하며 MBMT 또한 갱신 된다.

MBMT는 다음과 같은 정보를 지닌다.

- 원본 블록 주소  
원본 시스템 데이터의 시작 블록 주소
- 사본 블록 주소  
사본 시스템 데이터의 시작 블록 주소
- 길이  
대상 블록의 길이

시스템 데이터 영역이 손상되어 읽기 요청에 문제가 생겼을 경우, 대체 블록을 요청의 결과 값으로 돌려줘야 한다. 이를 위하여 대체 블록의 주소를 알 수 있어야 하며 이는 MBMT에 기록되어 있다. 하지만 MBMT는 모든 블록 주소를 기록하지 않으며 연속된 블록의 시작 블록의 주소만 기록하기 때문에 대상 블록이 소속된 연속 블록의 시작 블록을 탐색할 수 있어야 한다.

이러한 문제를 해결하기 위해 MBMT는 그림 10과 같은 B-트리로 구성된다. 그림 10은 동적 미러링 기법에 사용되는 B-트리의 예제로 MBMT의 원본 블록 주소를 기준으로 B-트리를 작성하였다. 대체 블록의 검색은 손상 블록과 연속적이며 MBMT에 기재된 시작 블록을 먼저 검색하여야 한다. 그림 11은 그림 9와 표 3의 데이터를 기반으로 작성한 검색 예제이다. 대체 블록 검색은 그림 11과 같이 B-트리를 순회하며 손상 블록의 시작 블록을 검색하여 대체 블록을 찾을 수 있다.

표 3. 미러 블록 매핑 테이블의 예제  
 Table 3. An Example of Mirror Block Mapping Table.

원본 블록 주소	미러 블록 주소	길이
0x00545070	0x00087154	25
0x00548370	0x00173145	27
0x01422754	0x04271726	33
0x01435584	0x05473218	14
...	...	...

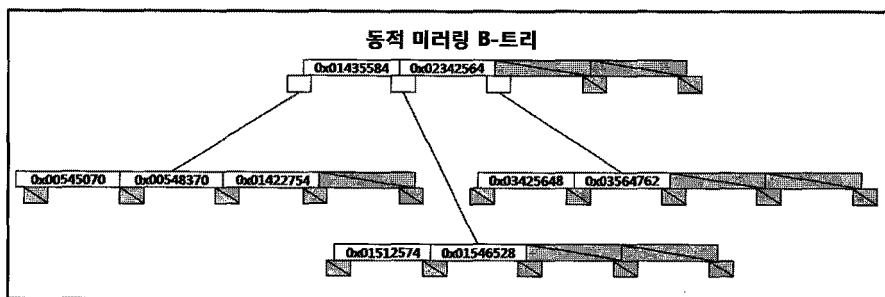


그림 10. 동적 미러링 B-트리의 예제  
 Fig. 10. An example of Dynamic Mirroring B-tree.

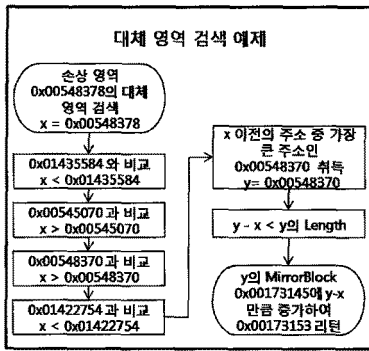


그림 11. 동적 미러링 B-트리의 검색 알고리즘 예제  
 Fig. 11. An example of searching algorithm of Dynamic Mirroring B-tree.

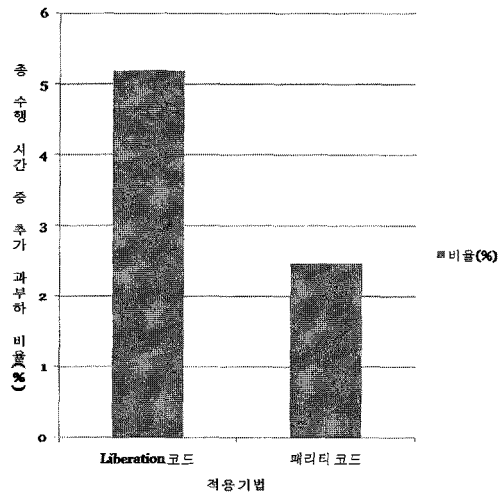


그림 12. 패리티 코드와 Liberation 코드의 Encoding 시간 비율 측정  
 Fig. 12. Time rate of encoding operation when applied parity code and Liberation code.

### V. 평가

본 장에서는 제시한 RSPM 기법을 적용한 파일 시스템과 기존의 파일 시스템의 성능을 비교한다. 그리고 채널 코딩 및 미러링 기법의 적용이 신뢰성에 어떠한 영향을 주는지를 평가한다.

RSPM 기법은 실험을 위해 리눅스 커널 2.6.24의 원시 코드를 일부 수정하였으며, EXT3 파일 시스템을 변형하여 RSPM 파일 시스템을 구축하였다. 또한 IP카메라에서 전송되는 카메라 데이터를 수신하는 비디오 서버 시뮬레이터를 작성하여 실험을 수행하였다<sup>[8-9]</sup>.

#### 1. 성능

제안하는 기법은 부호어를 저장하는 패리티 디스크로 하나의 플래시 메모리를 활용한다. 하지만 RSPM 기법에 적용된 Liberation 코드는 원래 2개의 패리티 디스크를 상정하여 작성되었으며 이로 인해 시스템 전체에 불필요한 과부하를 일으킬 수 있다. 표 4는 부호어 P만 생성하는 패리티 코드와 부호어 P, Q를 모두 생성하는 Liberation 코드의 XOR 연산 회수를 비교한 것이다. 단일 패리티 코드의 XOR 연산 회수를 1로 볼 때, Liberation 코드는 부호어 P, Q를 모두 생산하기 때문

표 4. 단일 패리티 코드와 Liberation 코드 XOR 연산 비율

Table 4. XOR operation rate of Parity code and Liberation code.

	인코딩	디코딩
패리티 코드	1.0	1.0
Liberation 코드	2.225	2.54

에 단일 패리티 코드의 2.2배 정도의 연산이 더 필요하다. 또한 약 2.54배의 연산이 디코딩 시 필요하게 된다.

Liberation 코드와 패리티 코드가 NVR 시스템의 성능에 미치는 영향을 측정하기 위해 다음과 같이 실험을 수행하였다. 4대의 IP 카메라에서 초당 24프레임으로 60분간 촬영하고 이를 네트워크로 전송받아 비디오 서버 시뮬레이터를 이용하여 저장하였다. 60분간 촬영 중 Liberation 코드는 약 187.013초의 인코딩 연산을 수행하였으며, 패리티 코드는 약 89.054 초의 인코딩 연산을 수행하였다. 그림 12는 60분의 촬영 시간을 100%로 가정할 때 인코딩 시간의 비율을 나타내며 Liberation 코드가 약 5.195%, 패리티 코드는 약 2.474%의 추가적인 과부하가 발생하였다. Liberation 코드는 고유의 코드 Q와 함께 패리티 코드 P도 함께 생성하므로 패리티 코드에 비하여 약 2배의 과부하가 발생하였다. 하지만 인코딩 연산은 시스템 전체의 연산 시간 중 낮은 비율을 차지하고 있으며 Liberation 코드를 적용 시 보다 높은 신뢰성을 확보할 수 있었다. 또한 실험 시 부호어 생성에 따른 추가적인 과부하는 느껴지지 않았으며 비디오 서버 시뮬레이터의 처리능력으로 충분히 분산하여 처리 가능한 정도였다. 따라서 본 기법의 적용에 따른 추가적인 과부하는 거의 없다고 볼 수 있다.

#### 2. 신뢰성

제안하는 기법은 파일 시스템 내의 데이터를 시스템



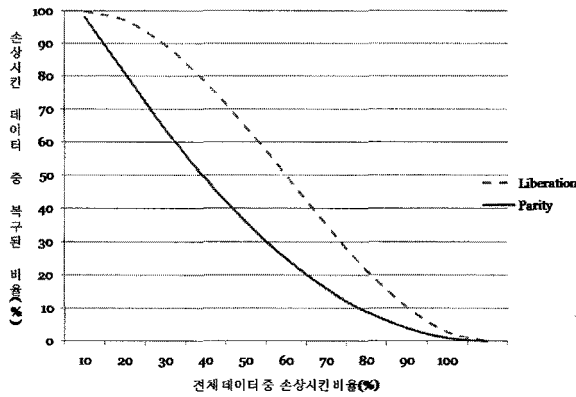


그림 13. 패리티 코드와 Liberation 코드로 난수 분포에 의해 손상된 데이터 복원율  
 Fig. 13. Recover rate of Parity code and Liberation code for damaged data by Random distribution.

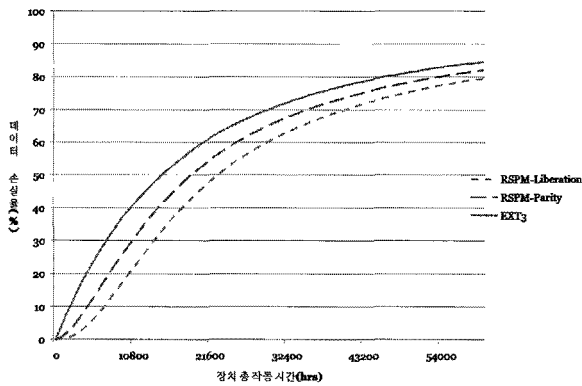


그림 14. 시간에 따라 와이블 분포에 의해 손상되는 데이터 중 복구되지 못한 데이터 손상률 ( $n = 9259, \beta = 1, k = 16$ )  
 Fig. 14. Recover rate of Parity code and Liberation code for damaged data by Weibull distribution.

타입과 미디어 타입, 두 가지로 분류하여 미디어 타입에 패리티 코드 및 Liberation 코드를, 시스템 타입에 동적 미러링 기법을 각각 적용하였다. k개의 데이터 블록이 주어질 때, Liberation 코드를 적용한 RSPM 기법은 2개의 손상된 블록을 복구할 수 있으며, 패리티 코드를 적용한 RSPM 기법은 1개의 손상된 블록만 복구할 수 있다. 따라서 패리티 코드가 Liberation 코드보다 연속적으로 일어나는 오류에 좀더 취약하다.

그림 13은 일종의 스트레스 테스트로 전체 하드디스크 중 생성된 오류의 비율과 이에 따른 패리티 코드와 Liberation 코드의 복원율을 보인다. 평균 약 15.177% Liberation 코드가 패리티 코드에 비하여 높은 복원율을 보였다.

그림 14는 와이블 분포에 따른 데이터의 손상률을 보인다<sup>[6]</sup>. 와이블 분포는 수명 데이터 분석에 주로 사용되는 연속형 확률 분포이다. 제안하는 기법은 불량 섹터와 같은 데이터 손실 오류에 대응하고자 하며 이에 대한 와이블 분포의 척도 인수는  $n = 9259$  으로 설정하였으며, 와이블 계수는  $\beta = 1$  로 설정하였다<sup>[4]</sup>.

그 결과 Liberation 코드를 적용한 RSPM 기법이 일반 EXT3 파일 시스템보다 평균 약 11.29%의 낮은 손상률을 보였으며 패리티 코드를 적용한 RSPM 기법보다 5.21% 낮은 손상률을 보였다.

### VI. 결 론

NVR은 IP 카메라와 비디오 서버를 네트워크로 구성하여 사용자에게 용이한 설치와 유지보수를 제공하지만 기존의 NVR 저장장치는 불량 섹터에 의해 발생될 수 있는 데이터 손실에 대한 대책이 부족하다. 본 논문에서는 NVR 저장장치에 저장되는 데이터를 미디어 타입과 시스템 타입의 두 가지 타입으로 분류하여 각각 적합한 신뢰성 정책을 적용하는 기법을 제안한다. 많은 공간을 필요로 하며 2차적인 복원 수단이 존재하는 미디어 타입은 패리티 코드와 Liberation 코드에 의한 부호어를 생성하며 이를 이용하여 데이터 손실을 복구한다. 제안하는 기법의 적용 결과 평균 약 11.289%의 데이터 손상을 방지할 수 있다. 적은 공간에 시스템의 실시간성을 유지하기 위해 필수적인 시스템 타입의 데이터는 손실 즉시 대체할 수 있는 동적 미러링 기법을 적용하여 예기치 않은 오류에도 실시간으로 촬영할 수 있는 환경을 제공할 수 있다.

### 참 고 문 헌

- [1] James S. Plank, "The RAID-6 Liberation Codes", Proceedings of the 6th USENIX Conference on File and Storage Technologies, Article No. 7, pp. 97-110
- [2] Kwanghee Park, Dong-Hwan Lee, Youngjoo Woo, Geunhyung Lee, Ju-Hong Lee, Deok-Hwan Kim, "Reliability and Performance Enhancement Technique for SSD array storage system using RAID mechanism", ISCIT 2009, pp. 140-145 Incheon, Korea
- [3] David A. Patterson, Garth Gibson, Randy H. Katz, "A case for redundant arrays of

inexpensive disks (RAID)", the 1988 ACM SIGMOD international conference on Management of data, pp 109-116, Chicago, Illinois, US, 1988.

[4] Elerath, J.G, Pecht, M., "Enhanced Reliability Modeling of RAID Storage Systems", Dependable Systems and Networks, DSN '07. 37th Annual IEEE/IFIP, pp. 175-184, 2007.

[5] Haryadi S. Gunawi, Vijayan Prabhakaran, Swetha Krishnan, Andrea C. Arpaci-Dusseau, Remzi H. Arpaci-Dusseau, "Improving File System Reliability with I/O Shepherdng", twenty-first ACM SIGOPS symposium on Operating systems principles, pp. 293-306, Stevenson, Washington, USA, 2007.

[6] John A. Chandy, Sumit Narayan, "Reliability Tradeoffs in Personal Storage Systems", ACM SIGOPS Operating Systems Review, Volume 41, Issue 1, pp. 37-41, 2007.

[7] Kwanghee Park, Jun-Sik Yang, Deok-Hwan Kim, Joon-Hyuk Chang, "Anticipatory I/O Management for Clustered Flash Translation Layer in NAND Flash Memory", ETRI Journal, Vol.30, No.6, pp. 790-798, 2008.

[8] Akshat Verma, Rohit Jain, Sugata Ghosal, "A Utility-Based Unified Disk Scheduling Framework for Shared Mixed-Media Services", ACM Transactions on Storage, Vol 3, No 4, pp. 17:1-17:30, 2008.

[9] 이근형, 박광희, 송재석, 김덕환, "비디오 서버 기반 저장장치의 신뢰성 향상을 위한 코딩 기법의 설계 및 구현", 2009 대한전자공학회 추계학술대회, 서울대학교, 대한민국, 2009.

[10] 김영진, 원유집, 김탁기, "SOAR : 저장장치를 기반으로 하는 시스템의 신뢰성 분석도구 개발", 정보과학회논문지: 시스템 및 이론, 제 35 제 6호, 2008.

[11] 우영주, 이근형, 박광희, 이동환, 김덕환, "EVENODD 코드 기법을 적용한 RAID-6 시스템의 성능 비교", 대한전자공학회 하계종합학술대회 제32권 제1호, pp. 607-608, 2009.

저 자 소 개



이 근 형(학생회원)  
 2008년 인하대학교 컴퓨터공학과  
 학사 졸업  
 2010년 인하대학교 전자공학과  
 석사 졸업.  
 <주관심분야 : 임베디드시스템,  
 스토리지 시스템>



김 덕 환(정회원)-교신저자  
 2003년 한국과학기술원 컴퓨터  
 공학 박사.  
 2006년~현재 인하대학교  
 전자공학부 부교수  
 <주관심분야 : 시각정보처리, 스  
 토리지 시스템, 임베디드 시스템>



송 재 석(학생회원)  
 2009년 대전대학교 전자공학과  
 학사 졸업.  
 2009년~현재 인하대학교  
 전자공학과 석사과정  
 <주관심분야 : 임베디드 시스템,  
 스토리지 시스템>